



Saridis, G. M., Peng, S., Yan, Y., Aguado, A., Guo, B., Arslan, M., Jackson, C., Miao, W., Calabretta, N., Agraz, F., Spadaro, S., Bernini, G., Ciulli, N., Zervas, G., Nejabati, R., & Simeonidou, D. (2015). LIGHTNESS: A deeply-programmable SDN-enabled data centre network with OCS/OPS multicast/unicast switch-over. In *European Conference on Optical Communication, ECOC* (Vol. 2015-November, pp. 1-3). [7341690] Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/ECOC.2015.7341690>

Peer reviewed version

Link to published version (if available):

[10.1109/ECOC.2015.7341690](https://doi.org/10.1109/ECOC.2015.7341690)

[Link to publication record in Explore Bristol Research](#)

PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via IEEE at <http://dx.doi.org/10.1109/ECOC.2015.7341690>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

LIGHTNESS: A Deeply-Programmable SDN-enabled Data Centre Network with OCS/OPS Multicast/Unicast Switch-over

G. M. Saridis⁽¹⁾, S. Peng⁽¹⁾, Y. Yan⁽¹⁾, A. Aguado⁽¹⁾, B. Guo⁽¹⁾, M. Arslan⁽¹⁾, C. Jackson⁽¹⁾,
W. Miao⁽²⁾, N. Calabretta⁽²⁾, F. Agraz⁽³⁾, S. Spadaro⁽³⁾, G. Bernini⁽⁴⁾, N. Ciulli⁽⁴⁾,
G. Zervas⁽¹⁾, R. Nejabati⁽¹⁾, D. Simeonidou⁽¹⁾

(¹) HPN, University of Bristol, U.K., (George.Saridis@bristol.ac.uk) (²) COBRA, Eindhoven University of Technology, the Netherland, (³) Universitat Politècnica de Catalunya, Spain, (⁴) Nextworks, Pisa, Italy

Abstract We demonstrate an all-optical dynamic DCN utilizing a fully SDN-enabled data-plane including programmable FPGA-based NICs, offering Network Function Virtualization OCS/OPS driven multicasting for Virtual DC applications enabling link recovery and VM migration.

Introduction

Next generation data centres are expected to deploy a network infrastructure that is agile, programmable and application adaptive. They will depend on ultra-high bandwidth and very low latency interconnection networks. To meet these expectations this paper proposes new technologies for both data and control planes of next generation Data Centre Networks (DCNs). In particular, we describe and demonstrate a virtualize-able and programmable hybrid all optical circuit and packet switched DCN infrastructure. Supporting multicast and hitless OCS/OPS switch-over it is managed by a Software Defined Network (SDN) controller empowered by monitoring Network Function Virtualization (NFV). This is the first demonstration of an all optical DCN with the capability to create multiple co-existing Virtual Data Centres (VDCs) enabled by SDN, where each VDC utilises NFV to deploy its own virtualised control functions from a repository of existing network functions. Furthermore, it is the first demonstration of NFV and SDN functionalities (monitoring and database migration) fully integrated with physical layer optical packet and circuit switched multicasting.

Overall architecture

The overall DCN architecture including both the optical data plane and control plane is illustrated in Fig. 1. Servers within the same cluster are interconnected via a novel optical Network Interface Card (NIC) and an all-optical top of the rack (ToR) switch to the programmable DCN¹. The NICs employ SDN-enabled hybrid OCS/OPS interfaces that support programmable composition and transmission of optical packets with associated labels or Ethernet frames. ToR and wavelength selective switches (WSS) are also SDN-enabled. The architecture includes SDN-enabled Optical Packet Switch (OPS) nodes with multicasting and optical packet reception/contention monitoring capabilities. In the proposed architecture, all OPS nodes and ToR switches of a single cluster are connected to a high-radix SDN-enabled optical backplane with OCS/OPS multicasting capability. The backplane hosts optical function block plug-ins (i.e. WSS, OPS, splitters for OCS multiplexing) and offers ToR-to-ToR and Cluster-to-Cluster connectivity. As such, it enables the SDN-based DCN control plane to build and reconfigure the data plane topology, by dynamically configuring

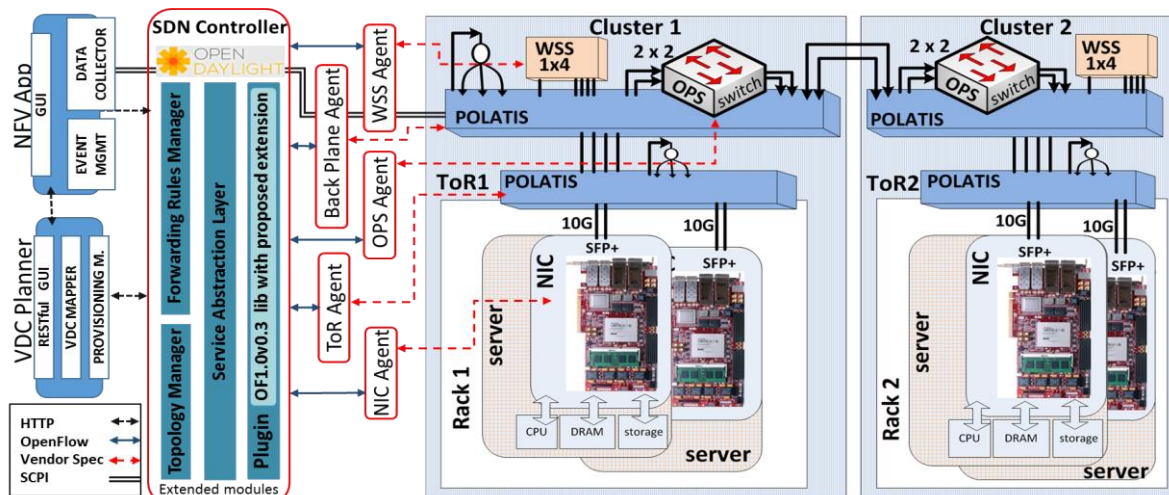


Fig. 1: Overall Data Centre Network architecture, control plane and experimental data plane

appropriate cross-connections in the optical backplane according to different applications' requirements. The FPGA-based OPS/OCS hybrid NIC² eliminates the electronic ToR switch and interconnects the intra-rack servers with ultra-low latency. The functionality of the novel NIC includes network interface functions, programmable aggregation and segregation functions, OPS/OCS switch and layer 2 switch functions. Thus, based on the DCN requirements and data flows, the FPGA-based OPS/OCS NIC can be configured by the SDN controller on-demand.

SDN-enabled control plane

For this experiment, OpenDaylight (ODL) is used as the SDN controller, and OpenFlow agents for Polatis, WSS, OPS and hybrid NIC were developed to enable further SDN-based programmability. In addition to the OpenFlow (OF) extensions³, shown in Fig. 1 left, OF is extended to allow configuration of the OPS packet duration in the NIC. Also, ODL internal modules are extended to support the following network device specific features. For the OPS ports and WSS, the Switch Manager and Service Abstract Layer (SAL) were extended to record the supported wavelength and supported spectrum range respectively, which are used to validate the configuration. In addition, the transmitted optical packet statistics can be recorded by the Statistics Manager. To properly configure these devices, the Forwarding Rules Manager was extended to construct the required configuration information e.g. *label & out-put* for the OPS, *central_freq & bandwidth & out-put* for the WSS and *match & label* (optional for OPS) & *out-put*.

The FPGA-based optical NIC communicates with the OF agent through a bidirectional

10Gbps SFP+ Ethernet interface. The commands and information are encapsulated in a 1504 Byte Ethernet frame. Furthermore, through the extended ODL, various applications communicate with the data-plane using the RESTful interface (see next section).

Virtual Data Centre (VDC) application

For the purposes of these tests, two control plane applications that sit on top of ODL have been deployed: a virtual data centre planner (VDC Planner) and a monitoring virtual network function (VNF). The user/VDC management deploys the VDC planner to dynamically create a VDC request. The parameters that can be specified are: *i*) servers to be used *ii*) links to be created *iii*) OCS or OPS for each link *iv*) multicast properties for servers and links. According to the request, the VDC planner generates flows to be programmed into the network, distributed among the different technologies (NIC, OPS, OCS and backplane). When a VDC has been deployed (in our case, a multicast VDC using OPS resources), the user can request a dynamic VNF to monitor various VDC parameters. The request is handled by a NFV server that creates the monitoring function. The created VNF retrieves network information via two different interfaces: RESTful northbound to the ODL controller to retrieve OPS packet counting information; and by using the backplane (Polatis power monitoring) to monitor the power of every link and device attached. Should degradation or failure be detected we can initiate a switch-over. The above information is compared with a theoretical maximum and a configurable threshold. The application enables two reactive scenarios: (i) when the optical power detected drops below a threshold value, switch the multicast traffic to a second OPS

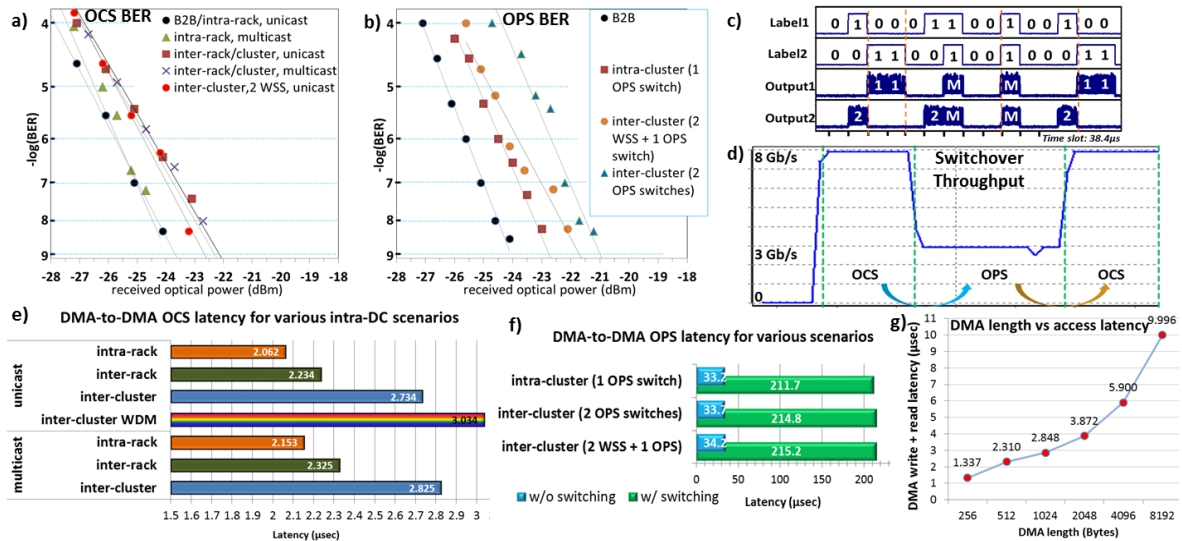


Fig. 2: a,b) BER curves, c) Time traces, d) Switchover throughput plot, e,f) DMA-to-DMA latency, g) DMA access latency

switch, and (ii) whenever the OPS packet rate exceeds the maximum, switch the multicast traffic using the optical splitter (OCS multicasting).

Experimental Demonstration & Results

The experiment is composed of four rack-mounted PowerEdge T630 servers each equipped with an FPGA-based NIC board utilizing 10G SFP+ transceivers, providing an interface to the optical network. The FPGA-based OPS/OCS hybrid NIC² using NETFPGA SUME development board, has been designed to plug directly into a server, and replace the traditional NIC. In the prototype design, it has an 8-lane Gen3 PCIe interface for DRAM communication, one 10Gbps interface for getting commands from SDN control agent and sending feedback, two OPS/OCS hybrid 10Gbps ports for inter-rack/cluster communication and an OPS label pin interface connecting to the OPS label generator.

All servers are connected to a 192×192 port Polatis circuit switch, which acts as a ToR switch and as an optical backplane on top of each cluster. A 1×4 optical power splitter, two 1×4 Wavelength Selective Switches (WSS) and one SOA-based 4×4 OPS⁴, which can logically perform as two 2×2 switches, are attached to that optical backplane. The splitter is used to accomplish OCS broadcasting scenarios; the WSSs for grooming inter-cluster channels from different servers/rack to the destination cluster. The OPS switch is able to rapidly switch optical packets with a reconfiguration time of 20ns. The label bits are generated by each NIC, then an optical RF tone label is created by the label generator and attached to each optical packet. At the OPS, the label is extracted processed and matched with the look-up table by the switch controller to determine the packets destination. Multicasting is enabled when two label bits have been set as “11” as shown in the time traces of Fig. 2c.

We measured BER for intra-rack, intra/inter-cluster scenarios, with one, two & without OPS switches using real traffic with PRBS payload. Minor penalties, <2dB, were observed for OCS (Fig. 2a) both for unicast and multicast, while OPS (Fig. 2b) shows 1 and <3dB penalties when passing through one and two switches respectively. Fig. 2d depicts the fluctuations of interconnection throughput when a switch to or from the OCS to OPS is initiated by the NFV and performed in the data plane. DMA-to-DMA latency measurements for different interconnection scenarios with OCS and OPS can be found in Fig. 2e & 2f. OPS latency w/ switching (Fig. 2f) includes considerable delays from

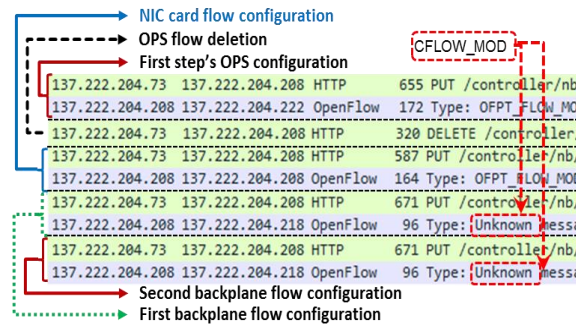


Fig. 3: Wireshark capture from SDN-enabled control plane

segregation, aggregation, the buffering and clock recovery functions from each of Tx & Rx FPGAs. Thus it is higher than OCS and OPS w/o switching. Actual DMA access latency was also measured (Fig. 2g) for different DMA lengths ranging from 256 to 8192 Bytes.

Fig. 3 shows the message exchanges among the VDC planner, ODL and the OF agents. The first two requests are performed to create the initial OPS-based VDC (only OPS static flow shown). The second set of messages reconfigures the network to go through the optical splitter (OCS multicasting), which consists of the OPS flow deletion, the creation of two cross-connections in Polatis and one change in the Tx NIC. Any set of requests consists of HTTP requests to the controller from the VDC planner & OpenFlow FLOW_MOD and CFLOW_MOD sequence exchange messages between the controller and the different agents.

Conclusions

This paper demonstrates for the first time a function and data-plane programmable all-optical DCN with a realistic migration scenario. The novel technologies demonstrated in this paper include: i) SDN-enabled and virtualizable programmable data plane, ii) SDN controller, NFV server supporting data plane monitoring and VM migration function virtualization, and DC virtualization (VDC planner) constituting the control plane.

Acknowledgements

This work is supported by EC FP7 grant no. 318606, LIGHTNESS project.

References

- [1] S. Yan et al, "Archon: A Function Programmable Optical Interconnect Architecture for Transparent Intra and Inter Data Center SDM/TDM/WDM Networking" in *IEEE/OSA JLT*, v. **33**, n. 8, pp. 1586–1595, (2015).
- [2] Y. Yan et al., "FPGA-based Optical Programmable Switch and Interface Card for Disaggregated OPS/OCS Data Centre Networks", *Proc. ECOC* (2015).
- [3] B. Guo et al., "SDN-enabled programmable optical packet/circuit switched intra data centre network" *Proc. OFC, Th4G.5*, (2015).
- [4] W. Miao et al., "SDN-enabled OPS with QoS guarantee for reconfigurable virtual data center networks," in *IEEE/OSA JOCN*, v. **7**, n.7, pp.634-643, (2015).